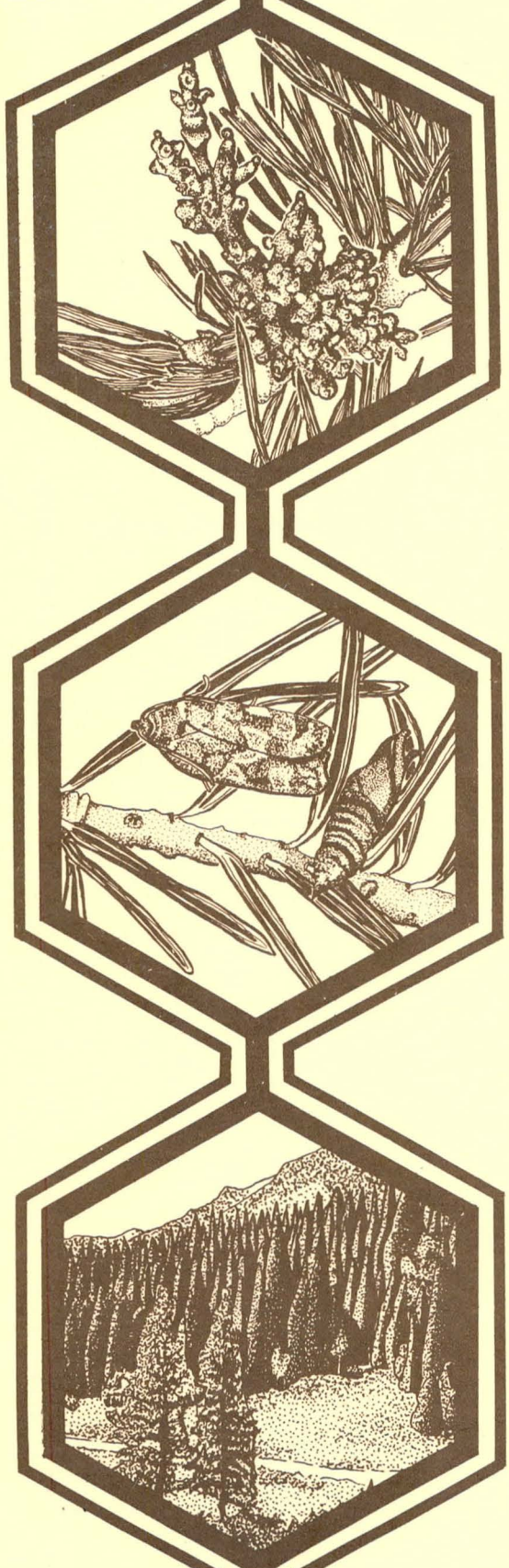
SB761
U8
#80-7

MAG Report 80-7



AN ANALYTICAL APPROACH FOR DEFINING OPTIMUM SAMPLE SIZE FOR SPRAY DEPOSIT ASSESSMENT

Methods Application Group
Forest Insect and Disease Management
Forest Service, USDA
2810 Chiles Road
Davis, California 95616

Report 80-7

3400
July 1980

AN ANALYTICAL APPROACH FOR DEFINING OPTIMUM SAMPLE SIZE FOR SPRAY DEPOSIT ASSESSMENT

John Wong[1]

ABSTRACT

A technique for defining optimum sample size procedures for spray deposit assessment was developed using an analytical approach. The technique was applied to data from a spray block from a pilot control project. Results derived are aimed at providing a decision-making tool to assist the pesticide specialist in planning a control project. Extension of the technique developed and the results derived for other conditions are possible under certain assumptions.

INTRODUCTION

In forest spraying the most commonly used indicators of application quality are spray droplet density, volume recovery rate, and volume median diameter (VMD)[2]. These performance indicators are usually expressed as drops/sq. cm, gallons/acre, and micrometers, respectively.

In most pilot and operational control projects, deposit assessment requires placement of four cards/tree in the cardinal directions directly under the drip line. The total number of spray deposit samplers required for each spray block ranges from 50 to 300, depending on the size of the area (Barry et al. 1979). In order to determine the degree of association between one or more of the performance indicators mentioned above and the change in insect population density, clusters of sample trees are established within a spray block. These clusters are sampling units to which standard regression analysis can be applied. For broader evaluation, i.e., for the entire spray area, data for the block level are used.

When conducting a control project over large areas, the cost of spray deposit samplers, the manpower required to place samplers in the field, subsequent processing costs, and, perhaps more importantly, environmental considerations, are major factors influencing magnitude of the

---

1 Mathematical Statistician, USDA-Forest Service, Forest Insect and Disease Management, Methods Application Group, Davis, CA.

2 The theoretical center of the spray droplet size spectrum at which half of the spray mass is in droplets larger than the VMD and half of the droplets smaller than the VMD.

deposit sampling effort. Recognizing this, the effect of altering the number of samplers must be known.

It is possible, by using historical field data, to evaluate the effect of reduced sample size on the resulting values of the performance indicators. These analyses can be made for the block and cluster level. Although these results have been derived for a specific data set, it will be demonstrated that under certain assumptions, the results are applicable to many conditions.

METHODS

Deposit data from a single block[3] of a pilot project conducted in 1976 by USDA Forest Service, Northern Region (R-1), to evaluate trichlorfon and acephate for managing the western spruce budworm, Choristoneura occidentalis Freeman, were used (Flavell et al. 1976). There were 25 clusters in this spray block, each consisting of three trees with four deposit samplers per tree, for a total of 300 deposit cards. A computerized analysis program (Young et al. 1977) which provides statistics on droplet density (D), mass recovery rate (R), and VMD (V) was used for our evaluations.

A random selection[4] of one, two, or three cards was made from each tree using a uniform probability density function in the interval (0,1) corresponding to each of these three alternatives to be compared relative to four cards. The selection process was repeated until the entire list of 75 trees was exhausted. Furthermore, 10 replicates or data sets were independently selected to represent the population. This population provides us with a means of testing and analyzing various sampling strategies with respect to the spray block and clusters within a spray block.

Comparison of these alternative with the four-card sample is based on the assumption that four cards provide an absolute measure of application quality and that operationally, the placement of more than four cards/tree is considered impractical.

---

3 The selected spray block was block number 3, which was sprayed with Orthene 75-S.

4 This method of selection is equivalent to random placement of cards under a tree.

RESULTS

Analyses of Block Means

For each alternative, the block means, i.e., $\bar{D}$, $\bar{R}$, and $\bar{V}$, were computed for each replicate. The following transformation equations were then applied to the results:

$$d_i = \frac{\bar{D}_i - \mu D}{\mu D} \times 100(\%)$$

$$r_i = \frac{\bar{R}_i - \mu R}{\mu R} \times 100(\%)$$

$$v_i = \frac{\bar{V}_i - \mu V}{\mu V} \times 100(\%)$$

where:

$\bar{D}_i$, $\bar{R}_i$, and $\bar{V}_i$ are the individual block means,

μ_D, μ_R, and μ_V are block means derived from four cards/tree,

d_i, r_i, and v_i are percentage deviations.

Since the distribution of $\bar{D}_i$, $\bar{R}_i$, and $\bar{V}_i$ is approximately normal (Goldstein 1965), it follows then that linear transformations of these variables also yield normal distributions. The resulting distributions (table 1) are the distributions of percentage differences of the output variables relative to the four-card configuration.

Confidence intervals for $\alpha = 0.05$ and $\alpha = 0.10$ were computed using the t distributions from the sample statistics above (table 2).

For any α, confidence intervals can be computed to provide a means of comparing various sampling intensities. For example, if the requirement is such that the droplet density for the selected alternative may deviate from the four-card configuration no more than $\pm$ 5 percent at $\alpha = 0.05$, then either two or three cards per tree will provide satisfactory estimates of deposit. If we assume that the distributions of differences hold for conditions different from those which occurred in this particular spray block, the derived results can provide the pesticide applicator a valuable tool for deciding on the intensity of deposit assessment sampling.

Table 1. Distribution of percent difference in block means for three spray deposit parameters relative to four cards/tree.

Replicate	One Card/Tree (Alternative 1)			Two Cards/Tree (Alternative 2)			Three Cards/Tree (Alternative 3)		
	$\bar{d}_i$	$\bar{r}_i$	$\bar{v}_i$	$\bar{d}_i$	$\bar{r}_i$	$\bar{v}_i$	$\bar{d}_i$	$\bar{r}_i$	$\bar{v}_i$
1	1.0	2.6	0.3	-2.4	-2.6	0.8	-0.4	-2.6	-0.1
2	-5.3	-7.9	0.3	1.8	2.6	1.5	1.7	0.0	-0.1
3	-1.6	-13.2	-4.7	-0.4	0.0	1.7	0.5	2.6	1.4
4	1.3	-2.6	-3.5	-2.4	0.0	3.2	-0.5	0.0	1.2
5	0.4	-5.3	-2.9	-0.7	0.9	-1.4	-0.2	0.0	1.0
6	4.3	7.9	1.0	-1.4	0.0	1.1	-1.4	-5.3	-0.2
7	-3.0	-5.3	-1.8	2.1	-2.6	-1.7	1.0	0.0	0.5
8	-1.2	0.0	-0.2	1.0	0.0	0.5	0.4	-2.6	0.1
9	3.8	0.0	-4.4	0.3	2.6	0.3	-1.3	-2.6	1.6
10	1.1	0.0	-0.7	-1.1	-2.6	0.2	-0.4	-2.6	0.3
Mean	0.08	-2.38	-1.66	-0.32	-0.17	0.62	-0.06	-1.31	0.57
S.E.[5]	2.94	5.89	2.09	1.60	1.95	1.44	0.97	2.23	0.68

Table 2. Confidence intervals for percentage difference in block means relative to the four-card configuration.

Parameters Evaluated[5]	95%			90%		
	Alt. 1	Alt. 2	Alt. 3	Alt. 1	Alt. 2	Alt. 3
$\bar{d}_i$	(-6.6,6.7)	(-3.9,3.3)	(-2.2,2.1)	(-5.3,5.5)	(-3.2,2.6)	(-1.8,1.7)
$\bar{r}_i$	(-15.7,10.9)	(-4.6,4.2)	(-6.3,3.7)	(-13.1,8.4)	(-3.7,3.4)	(-5.4,2.8)
$\bar{v}_i$	(-6.4,3.1)	(-2.6,3.9)	(-1.0,2.1)	(-5.5,2.2)	(-2.0,3.2)	(-0.7,1.8)

5 Standard deviation of the means from the 10 replicates.

Analysis of Clusters

The change in insect population density as a result of suppression projects is a key measure of treatment efficacy. This measure provides a means for comparing one or more insecticides. For evaluations, the cluster would then become the sampling unit for which insect density is observed. Spray droplet density is often correlated with insect mortality on a cluster basis. Distributions of cluster means would then become the major area of interest.

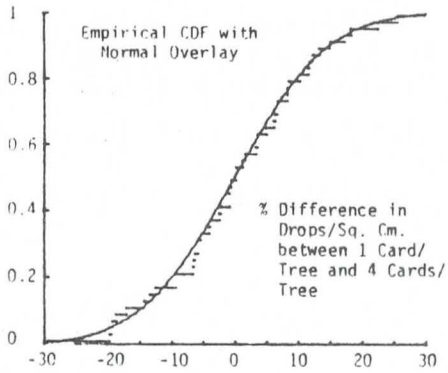
To evaluate these distributions, a random sample of five clusters from the same population were analyzed. Since each cluster contributes 10 cluster means from 10 replicates, our sample then consisted of 50 sampling units. For each cluster we assume that the distribution of cluster means from the replicates are approximately normally distributed. Our analysis focuses as before on the transformed variables $\bar{d}_j$, $\bar{r}_j$, and $\bar{v}_j$. The mean and standard deviation for each of the three alternatives are summarized in table 3.

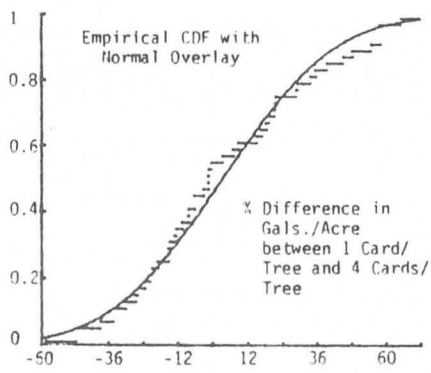
Table 3. Statistics for the distribution of percentage difference cluster means relative to four cards/tree.

Alternative Evaluated	$\bar{d}_j$ (%)		$\bar{r}_j$ (%)		$\bar{v}_j$ (%)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1	-0.07	11.80	3.28	30.91	0.96	10.17
2	0.03	6.50	1.17	14.80	0.86	6.34
3	-0.06	4.12	-0.90	10.70	0.14	3.33

In computing the statistics for each variable we assume that each sample is drawn from a population which is normally distributed. With this, the cumulative distribution functions (CDF's) are constructed as shown in figures 1a-1c, which can be used to compare various alternatives under a probabilistic framework for any cluster. In these figures, the cumulative probabilities are represented on the Y-axis and the range of percent differences are represented on the X-axis. The probability of occurrence between two points on the X-axis is arrived at by computing the difference between the two corresponding cumulative probabilities from the Y-axis. The continuous curves are the normal approximation to CDF's derived from empirical data. Each line segment on the empirical CDF curves represents those events which have the same probability of occurrence as derived from the frequency distribution. The normal approximation was a valid procedure based on the Kolomokorov-Smirnov test[6] at $\alpha = 0.05$ or 5 percent level. To

6 The one sample analysis program from a Tektronix 4051 microprocessor was used for this test.





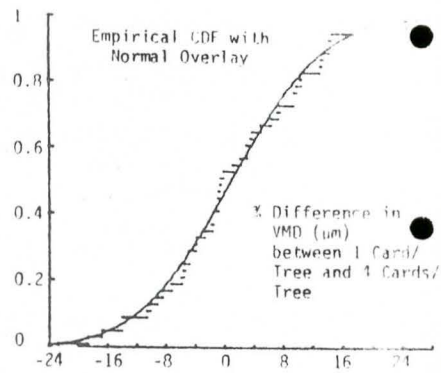
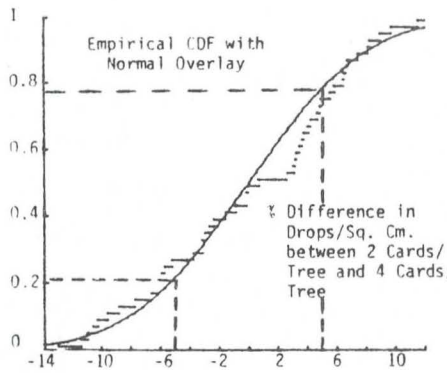
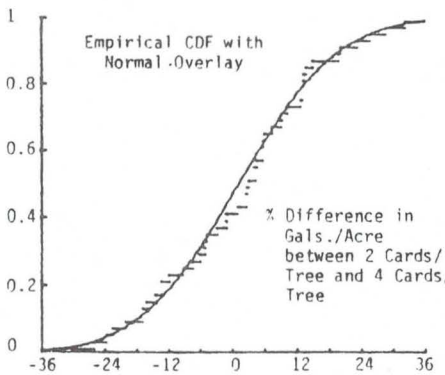


Figure 1a. Cumulative distribution functions comparing 1 card vs. 4 cards for clusters.





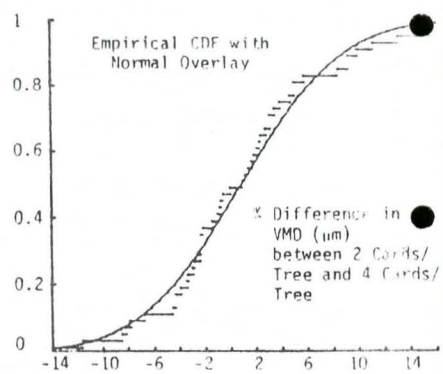
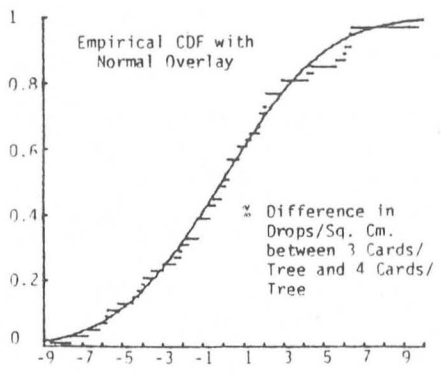
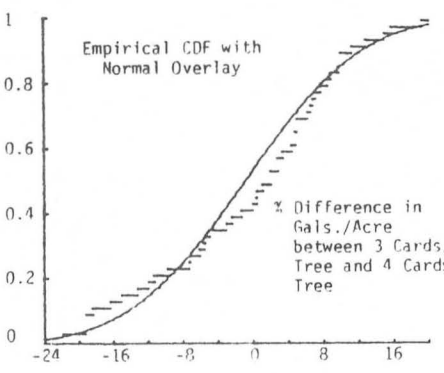


Figure 1b. Cumulative distribution functions comparing 2 cards vs. 4 cards for clusters.





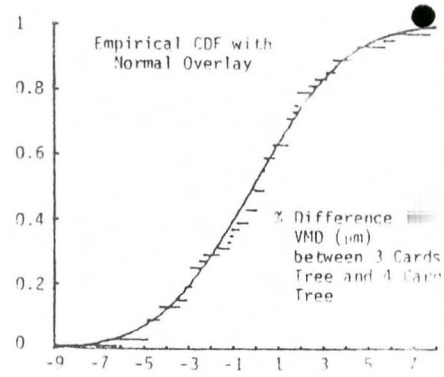


Figure 1c. Cumulative distribution functions comparing 3 cards vs. 4 cards for clusters.

Legend: Y-axis = cumulative probability
X-axis = percent difference

illustrate how one might use these results, assume that in planning a control project, a pesticide applicator decides that the percentage difference between the selected alternative relative to the four-card configuration on droplet density and recovery must not exceed $\pm$ 5 percent and $\pm$ 10 percent respectively with at least 50 percent probability. Referring to the normal CDF's, the probabilities of occurrences for these events are summarized in table 4. To arrive at the figure, for example, for drops/sq. cm. under alternative 2, i.e., 55.0 percent, as reported in table 4, the first graph in figure 1b is used. The $\pm$ 5 percent lower and upper limits for the amount of acceptable deviation are first located on the X-axis, the corresponding cumulative probabilities can then be determined. These probabilities are 0.22 and 0.77. The dotted lines are constructed to illustrate how this procedure works. The probability that the percent deviation in drops/sq. cm is within $\pm$ 5 percent is the difference between P = 0.77 ($X \leq 5$ percent) and P = 0.22 ($X \leq -5$ percent), or 55 percent. It is evident from examining these probabilities that the logical choice would be alternative 3, i.e., three cards/tree. In practical situations, the set of requirements can be considerably more complex.

Table 4. Probabilities of occurrence for individual clusters.

Parameters Evaluated	Alternatives		
	1	2	3
$\bar{d}_i$	23.0 (%)	55.0 (%)	76.0 (%)
$\bar{r}_i$	28.0 (%)	48.0 (%)	66.0 (%)

DISCUSSION

The results derived for the block and for the individual clusters are based on the assumption that these distributions are normally distributed. While it is feasible that a larger number of replicates can be generated, it is felt that the variance for $\bar{d}_i$, $\bar{r}_i$, and $\bar{v}_i$ for the block is sufficiently accounted for from the 10 replicates. The t-distribution was used to construct the confidence intervals in the analysis of block means to compensate for the fact that only 10 replicates were generated. In using the derived cumulative distribution functions for evaluating various alternatives at the cluster level, the choice to use either the empirical or the normal CDF should be made on the basis of regions of interests on these graphs, i.e., for a given value on the X-axis the associated probability can differ quite significantly by examination.

It remains to be shown that our results are valid for conditions other than those applicable to the selected spray block. Let us assume that this spray block was subjected to a different set of conditions, for example, the terrain, the meteorological conditions, the spray material, and the aircraft were different. Further, assume that with this new set of conditions, the effect on the performance variables are such that their magnitudes would have been changed by a multiplicative factor, $\pm$ k, where k is a real number (except zero) applied equally to either one, two, three, or four cards/tree. Consider the transformation equation used above for the analysis of block means on droplets. If μ_D is increased (or decreased) by a factor k which yields $k\mu_D$, then $\bar{D}_i$ becomes $k\bar{D}_i$, and therefore d_i remains unchanged. It follows then that the distribution of percent differences and the results derived from it are unaffected under these assumptions for the block and the individual clusters.

In planning a control project a pesticide applicator can now use the techniques developed and/or the results derived herein as a decision support aid with which cost components can be integrated. It is conceivable that with sufficient cost information, the technique of a marginal analysis can be applied to ascertain the amount of information gain in terms of the increased precision in our performance variables estimates relative to the increase in sampling intensity.

REFERENCES CITED

Barry, J.W., R.B. Ekblad, G.P. Markin, and G.C. Trostle (editors). 1978. Methods for sampling and assessing deposits of insecticidal sprays released over forests. USDA Tech. Bull. 1596. 162 pp.

Flavell, T.H., S. Tunnock, and H.E. Meyer. 1976. A pilot project evaluating trichlorfon and acephate for managing the western spruce budworm, Helena National Forest--1976. USDA-For. Serv., Northern Region, State and Private Forestry, Missoula, MT. Rpt. No. 77-16. 49 pp.

Goldstein, A. 1965. Biostatistics. The Macmillan Co., NY; Collier-Macmillan Ltd., London. 272 pp.

Young, R.W., R.C. Luebbe, and J.W. Barry. 1977. ASCAS--Data management system for assessment of aerial spray deposits. USDA-For. Serv., Forest Insect and Disease Mgmt., Methods Application Group, Davis, CA. Rpt. No. 77-2. 29 pp.